Coherent Coincident Analysis of LIGO Burst Candidates

Laura Cadonati
Massachusetts Institute of Technology
LIGO Scientific Collaboration

8th Gravitational Wave Data Analysis Workshop
Milwaukee, Wisconsin, December 17-20, 2003

LIGO-G030691-00-Z

Post Coincidence Coherent Analysis

- Burst candidates separately identified in the data stream of each interferometer by the Event Trigger Generators (ETG): TFclusters, Excess Power, WaveBurst, BlockNormal.
  - » Tuning maximizes detection efficiency for given classes of waveforms and a given false rate ~ 1-2 Hz
- Multi-interferometer coincidence analysis:
  - » Rule of thumb: detection efficiency in coincidence ~ product of efficiency at the single interferometers. Coincidence selection criteria should not further reduce the detection efficiency. The final false rate limits how loose the cuts can be.
  - » Currently implemented: time and frequency coincidence (in general, different tolerance for different trigger generators).
  - » Amplitude/energy cut: not yet implemented.
- Cross-Correlation for coherent analysis of coincident events
  - » This is a waveform consistency test.
  - » Allows suppression of false events without reducing the detection efficiency of the pipeline.

r-statistic Cross Correlation Test

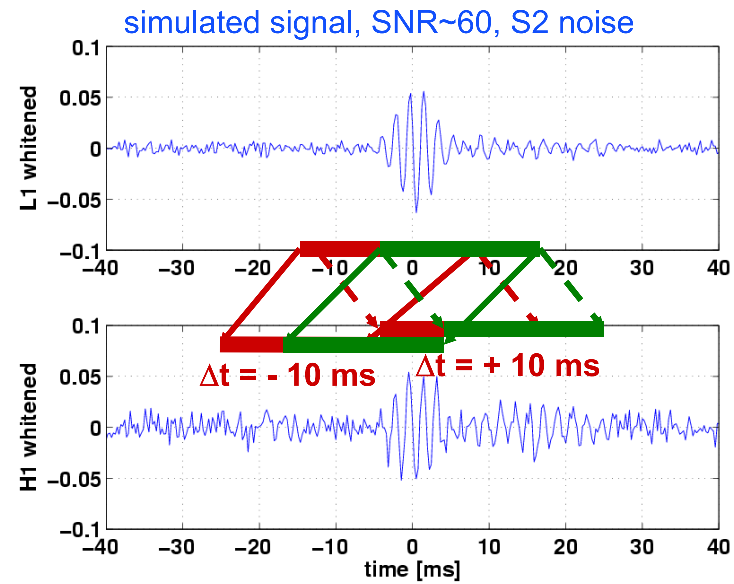
For each triple coincidence candidate event produced by the burst pipeline (start time, duration ΔT) process pairs of interferometers:

Data Conditioning:

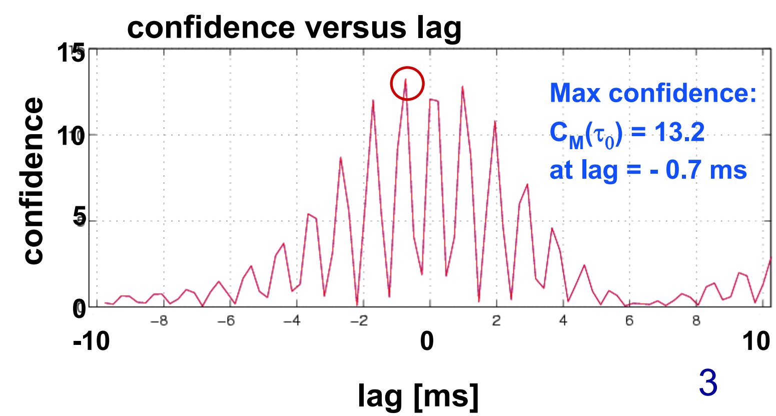
- 100-2048 Hz band-pass
- Whitening with linear error predictor filters

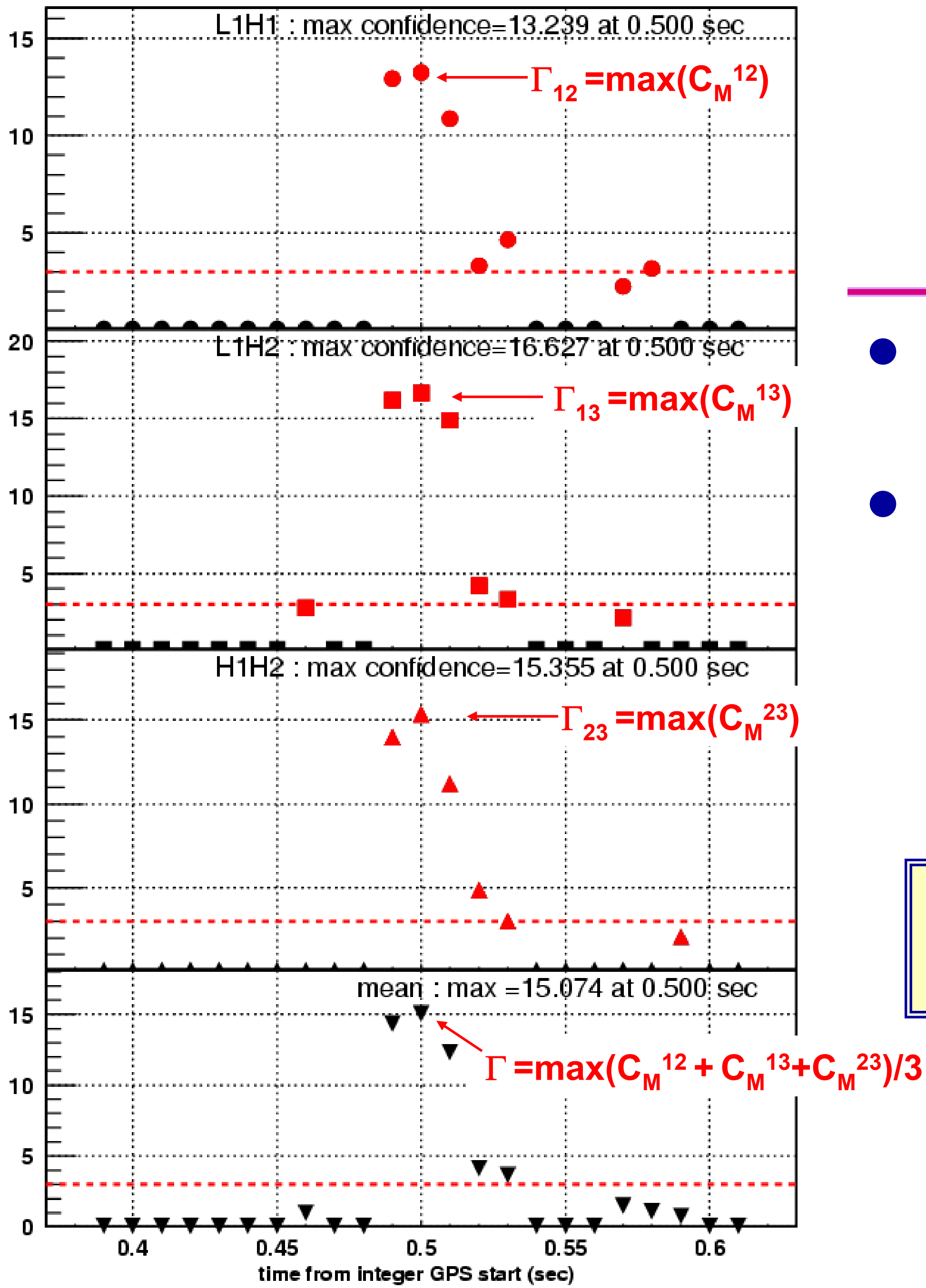
Partition the trigger in sub-intervals (50% overlap) of duration $\tau =$ integration window (20, 50, 100 ms).
For each sub-interval, time shift up to 10 ms and build an r-statistic series distribution.

$$r_k = \frac{\sum_i (x_i-\bar{x})(y_{i+k}-\bar{y})}{\sqrt{\sum_i (x_i-\bar{x})^2}\,\sqrt{\sum_i (y_{i+k}-\bar{y})^2}}$$



If the distribution of the r-statistic is inconsistent with the no-correlation hypothesis: find the time shift yielding maximum correlation confidence $C_M(j)$ (j=index for the sub-interval)





C_M(j) plots

- Each point: max confidence C_M(j) for an interval τ wide (here: τ = 20ms)
- Threshold on Γ:

  2 interferometers:

  Γ=max$_j$(C_M(j)) > β_2

  3 interferometers:

  Γ=max$_j$(C_M^{12}+ C_M^{13}+ C_M^{23})/3 > β_3

In general, we can have $\beta_2 \neq \beta_3$

β_3=3: 99.9% correlation probability in each sub-interval

Testing 3 integration windows:
20ms (Γ_{20}) 50ms (Γ_{50}) 100ms (Γ_{100})
in OR: Γ=max(Γ_{20},Γ_{50},Γ_{100})

Triple Coincidence Performance Analysis in S2

Exploring the test performance for triple coincidence detection, independently from trigger generators and from previous portions of the analysis pipeline:

- Add simulated events to real noise at random times in the 3 LIGO interferometers, covering 10% of the S2 dataset (*in LIGO jargon: triple coincidence playground*)
- apply r-statistic test to 200 ms around the simulation peak time

Definition of quantities used to characterize a burst signal:

$$h_{rss} = \sqrt{\int_0^{\infty} |h(t)|^2 dt} = \sqrt{\int_{-\infty}^{\infty} |\tilde{h}(f)|^2 df}$$

Total energy in the burst (units: strain/rtHz)
[directly comparable to sensitivity curves]

$$SNR = \sqrt{2 \int_0^{\infty} \frac{|\tilde{h}(f)|^2}{S_h(f)} df} \approx \frac{h_{rss}}{\sqrt{S_h(f_c)}}$$

SNR definition for excess-power techniques in the burst search = $SNR_{matched\ filtering} / \sqrt{2}$

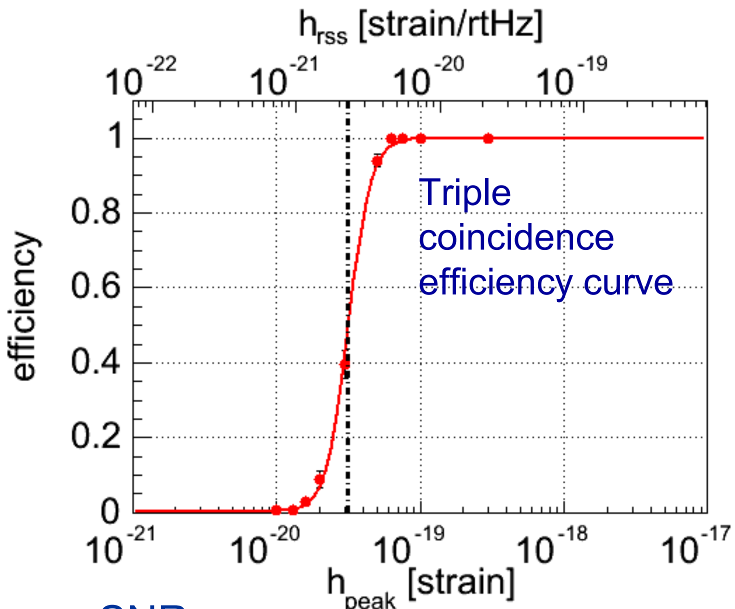
For narrow-band bursts with central frequency f_c

$S_h(f)$=single-sided reference noise in the S2 Science Run
⇒ reference S2 SNR for a given amplitude/waveform

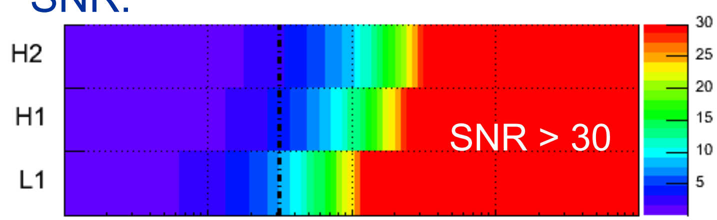
Detection Efficiency for Narrow-Band Bursts

Sine-Gaussian waveform f_0=254Hz Q=9
linear polarization, source at zenith

$$h(t) = h_{peak} \sin(2\pi f_0 (t-t_0))\, e^{-(t-t_0)^2/\tau^2} \qquad Q=\sqrt{2}\pi\tau f_0$$

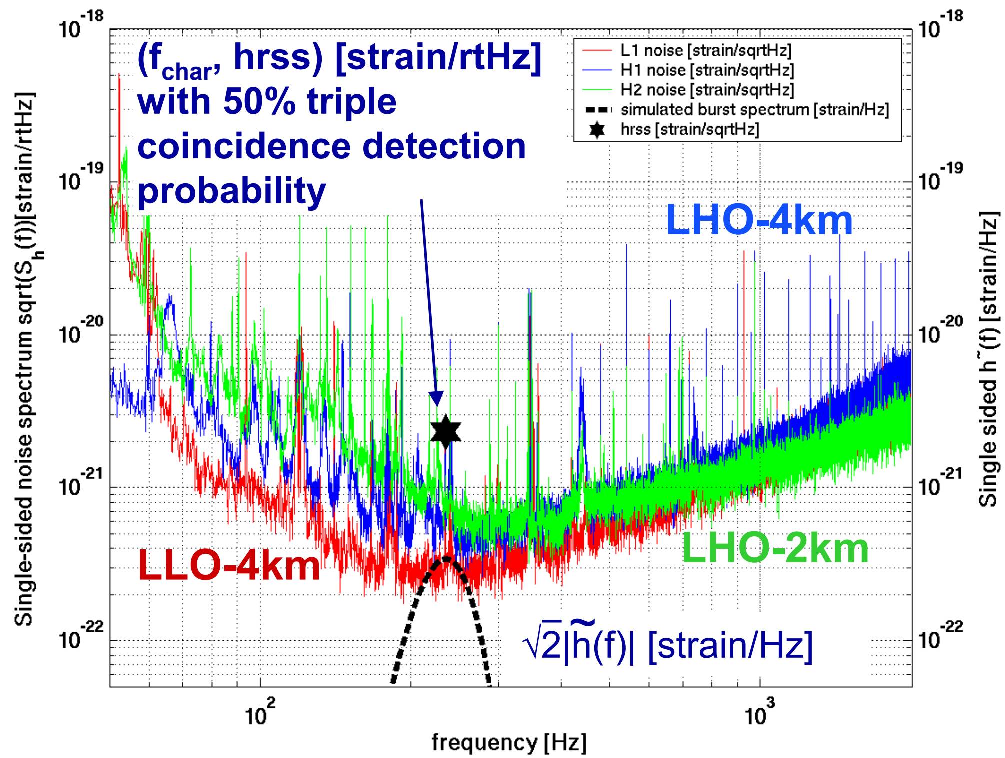


SNR:



$$h_{rss} = \sqrt{\int_0^\infty |h(t)|^2 dt} = \sqrt{\int_{-\infty}^{\infty} |\tilde{h}(f)|^2 df} = \sqrt{\frac{Q}{4\sqrt{\pi} f_0}}\, h_{peak}$$

$$SNR = \sqrt{2 \int_0^\infty \frac{|\tilde{h}(f)|^2}{S_h(f)} df} \approx \frac{h_{rss}}{\sqrt{S_h(f_c)}}$$



50% triple coincidence detection probability:

h_{peak} = 3.2e-20 [strain] h_{rss} = 2.3e-21 [strain/rtHz]

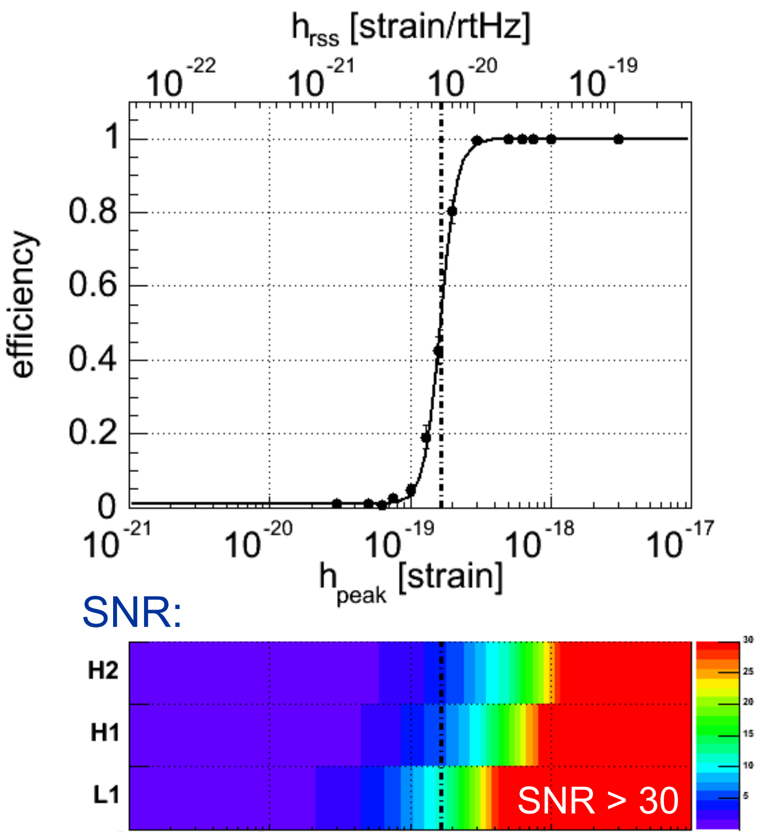
SNR: LLO-4km=8 LHO-4km=4 LHO-2km=3

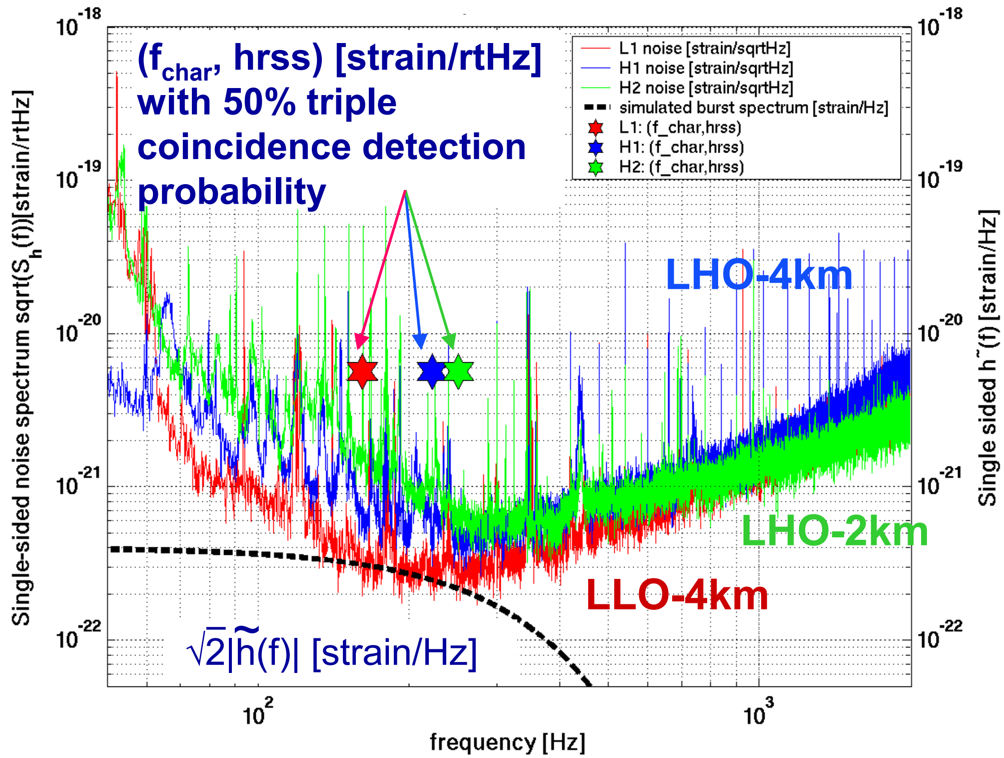
Detection Efficiency for Broad-Band Bursts

Gaussian waveform τ=1ms

linear polarization, source at zenith

$$h(t) = h_{peak}\, e^{-(t-t_0)^2/\tau^2}$$





50% triple coincidence detection probability:

h_{peak} = 1.6e-19 [strain] h_{rss} = 5.7e-21 [strain/rtHz]

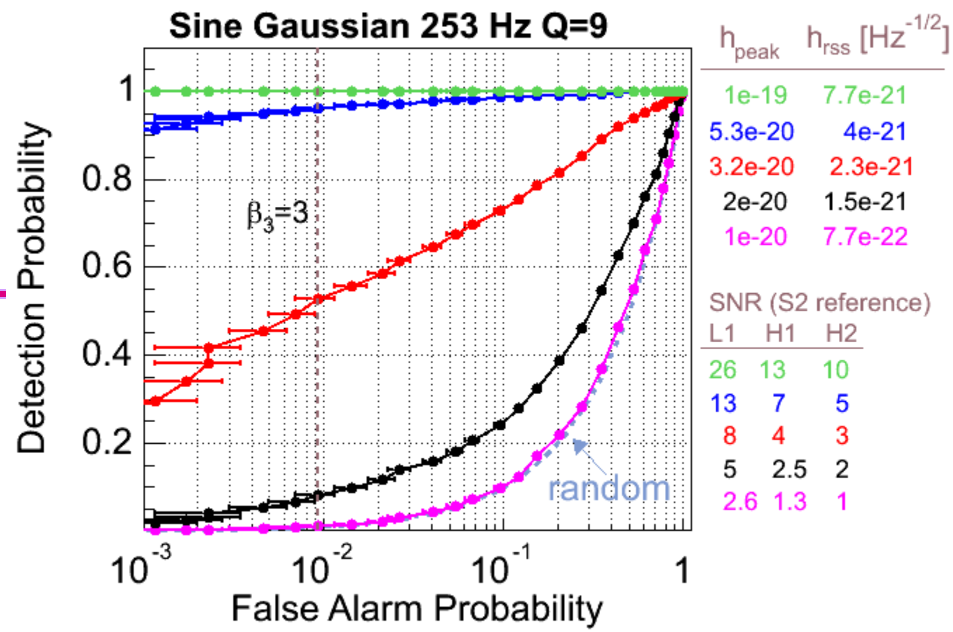
SNR: LLO-4km=11.5 LHO-4km=6 LHO-2km=5

$$SNR = \sqrt{2\int_0^\infty \frac{|\tilde{h}(f)|^2}{S_h(f)}\, df}$$

R.O.C.

Receiver-Operator Characteristics

Detection Probability versus False Alarm Probability.
Parameter: triple coincidence confidence threshold β_3

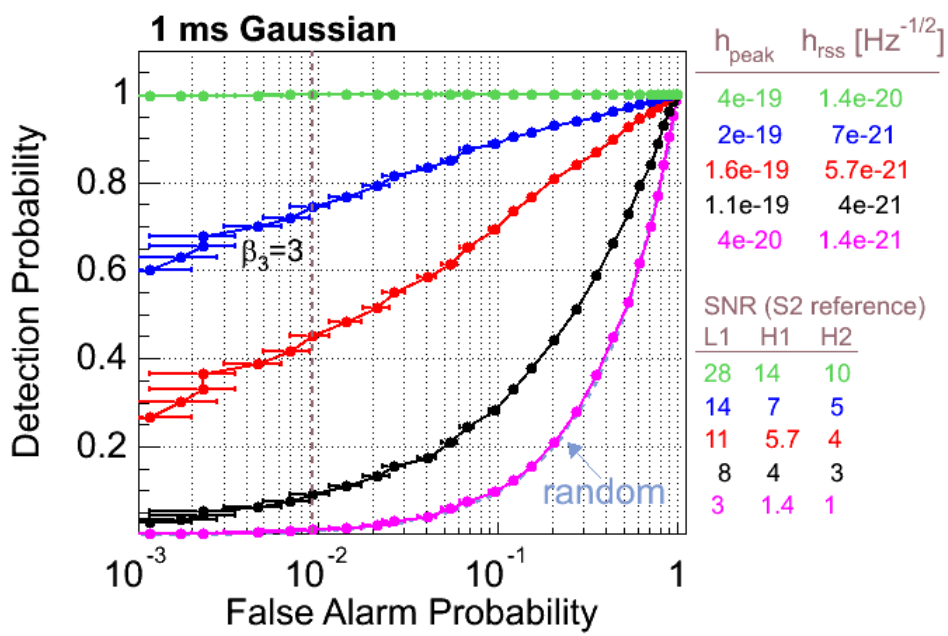


Simulated 1730 events at fixed h_{peak} ,h_{rss}
(10 events uniformly distributed in each S2 "playground" segment)

Tested cross correlation over 200 ms around the peak time

Operating condition: β_3=3

chosen from first principles (99.9% correlation probability in each event sub-interval for a pair of interferometers), corresponds to a ~1% false alarm probability for triple coincidence events with duration 200 ms.



Suppression of Accidental Coincidences from the Pipeline

In general: depends on the Event Trigger Generator and the nature of its triggers.
In particular: typical distribution of event duration (larger events have more integration windows).
Shown here: TFCLUSTERS 130 - 400 Hz
(presented in Sylvestre's talk)

Triple Coincidence Playground. T=88800 s (24.7 hours)	*Singles*
LLO-4km (L1)	2.5 Hz
LHO-4km (H1)	2 Hz
LHO-2km (H2)	2 Hz

Coincident numbers reported here are averages of 6 background measurements:
LLO-LHO = **± 8, ± 6, ± 4 sec** (H1-H2 together)

PRELIMINARY!!

"Loose" coincidence cuts

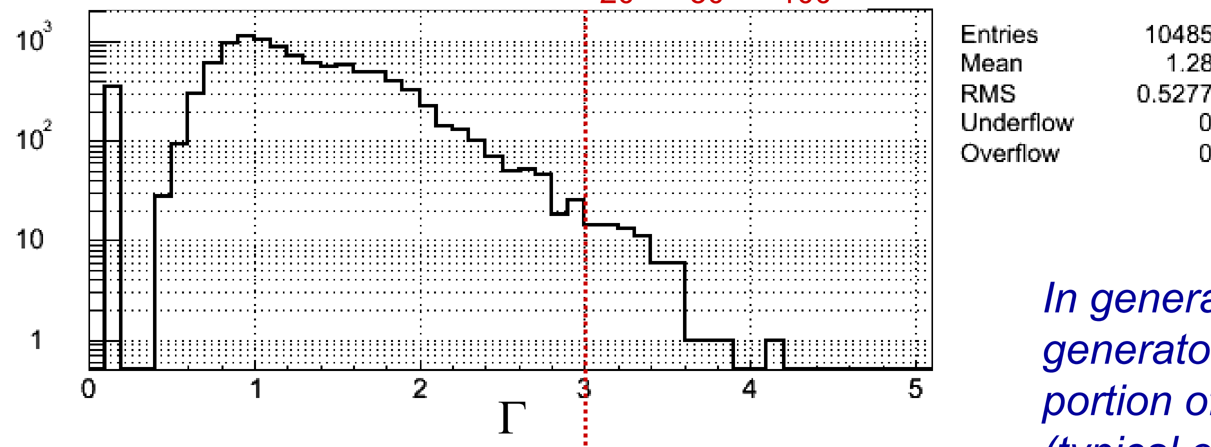
coincidence	*triple coincident clusters (Δt = 30 ms)*	*after frequency cut (200Hz tolerance)*	*after r-statistic test (β_3 = 3)*	***Rejection efficiency***
L1-H1-H2	20 mHz	15 mHz	0.1 mHz	(99.35 ± 0.08)%

"Tight" coincidence cuts

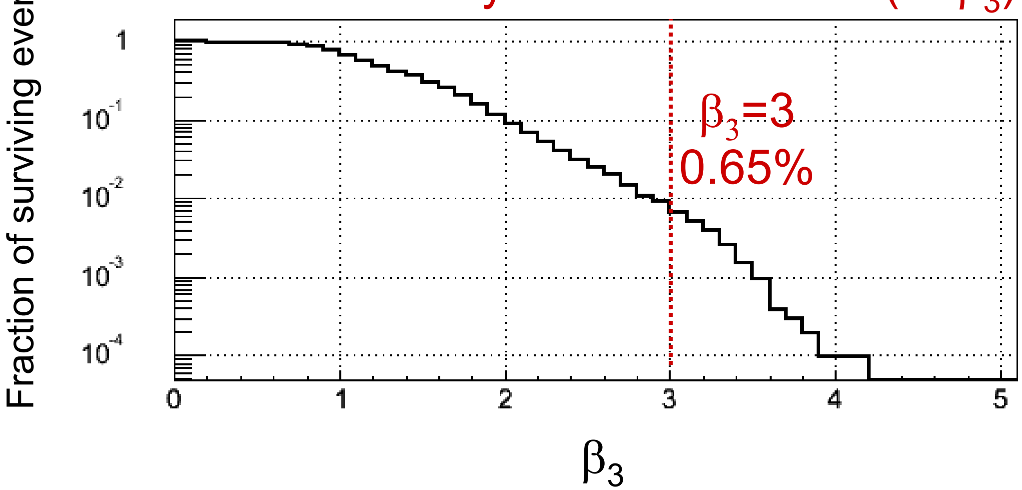
coincidence	*triple coincident clusters (Δt = 15 ms)*	*after frequency cut (75Hz tolerance)*	*after r-statistic test (β_3 = 3)*	***Rejection efficiency***
L1-H1-H2	6 mHz	1 mHz	0.01 mHz (1/day)	(98.8 ± 0.4)%

False Probability versus Threshold

Histogram of $\Gamma = \max(\Gamma_{20}, \Gamma_{50}, \Gamma_{100})$



False Probability versus threshold ($\Gamma > \beta_3$)



In general: depends on the trigger generators and the previous portion of the analysis pipeline (typical event duration, how stringent are the selection and coincidence cuts)

Shown here:
TFCLUSTERS 130-400 Hz with “loose” coincidence cuts

Conclusions

- The LIGO burst S1 analysis exclusively relied on event trigger generators and time/frequency coincidences.
- The search in the second science run (S2) includes a new module of coherent analysis, added at the end of the burst pipeline:
  r-statistic test for cross correlation in time domain
  - Assigns a confidence to coincidence events at the end of the burst pipeline
  - Verifies the waveforms are consistent
  - Suppresses false rate in the burst analysis, allowing lower thresholds
- Tests of the method, using simulated signals on top of real noise, yield 50% triple coincidence detection efficiency for narrow-band and broad-band bursts at SNR=3-5 in the least sensitive detector (LHO-2km) with a false probability ~1%.
- Currently measuring global efficiency and false rate for the S2 pipeline (event analysis + coherent analysis).